

SEQUENCE LISTING

<110> HAFEN, Ernst

<120> In vivo model system for type-2 diabetes

<130> 27656/38053

<140> To be determined

<141> To be determined

<150> PCT/IB99/01166

<151> 1999-06-22

<160> 4

<170> PatentIn Ver. 2.1

<210> 1

<211> 2907

<212> DNA

<213> Drosophila melanogaster

<400> 1

atggcatcaa tatcgatga cggcatggcg ctgagtggct acctcaagaa gctgaagacc 60
atgaagaaga agttctttgt gctgtacgag gagacgagca cttcggcagc ccggctggag 120
tactacgata ccgaaaagaa gttcctgcaa agagccgagc caaaaagggt tatatatctg 180
aagaattgct tcaacatcaa tcgccgtttg gacaccaagc atagatttgt cattgtgctc 240
tcctccagag acggtggatt cggcatcggt ctcgagaacg aaaatgattt acgcaaattg 300
ttggacaaac tactagtctt acaaaggaac atagccaatt cgaatggaac agcgactca 360
ccttatgacc acgtttggca agttgtcatt caaaagaagg gtatttcgga gaaagttgga 420
atcacccgaa cctaccactg ttgccttact tcaaaatccc tgacattcgt gtgcattgga 480
ccggagaaga cgcccaatgg cgaggatcgc gttgcgagca ttgaaatact tttgaccacg 540
atcaggcgat gcggtcatgc atccccacaa tgtatattct acgtggaact tggccgccaa 600
agtgtcttgg gatctggtga tctgtggatg gagacggata acgcagctat tgctactaat 660
atgcacaaca cgatactgag cgctatgtca gccaaaacag agtcgaacac gaatttaata 720
aacgtttatc agaatagacc tgacttaagt cagcagccca tgagaaagcg atcgtcgtct 780
gcaaacgaag catcgaagcc gataaacgta aatgtcatac aaaatagtca aaactctctc 840
gaattgcgca gctgcagttc gcccataac tatgggtttcg gcagagagag atgcgatagc 900
ttaccaacca gaaatggaac cctaagcgag tccagcaatc aaacgtactt tggttccaac 960
catggactgc gatccaatac tatatctggc atccgtccgc actcaaccaa caagcatagt 1020
aatagtccaa cgttcaccat gccattaaga tgctcagaat ccgaagagtc atcaattagt 1080
gtcgatgaat ccgacgacaa cggcagtttt agccactaca gattaaacac gcggtcatct 1140

gagacggcaa ttcttgagga aaacattgat gactttgccg gtgcggaatt atttagcaaa 1200
gtcaccgaac aaaatgtaag tgacgaaaac tacataccga tgaatccagt caatcctacc 1260
gatgctatcc atgaaaagga gaaggctgat atgcagagat tggaagatgc ttgctgcat 1320
ttcaactttc cggagcacgc gtcggaaaag cttgctaagg attttgatct ggactctgat 1380
aaccaatgct gtctgcccac tcgcgcctat tcgataggca acaaggttga gcatttaaag 1440
tttaataagc gcctgggaca cttgaatgat acgggacaga atccgaatcg cgtgcgagcc 1500
tactcggttg gctccaaatc gaagataccg cgctgcgacc tgcagcgagt ggctctctg 1560
gaggacaata aacatgagtt cacagcgaat aggagtcaga gtagcattac caaggaagga 1620
accagctatg gcagcagtcg caatcgacaa aagaagtcca caagtgtctc actcctcagt 1680
ctgaagaacc agataaaactc cgaccgaatg agtgacttaa tggaaattga tttttcacia 1740
gcaaccaatt tggaaaagca gaagttcatc aagaataatg aaattccgaa atacattgaa 1800
aacgtgttcc caaaagcccc gcgaacggat agctccagcc taactctgca cgccacaagt 1860
caaaaggaca ttttcaatgg caccaaacta aataaactg cgatcacatc cgaggatgg 1920
tacctcgaga tgaagccagt cggtaatgga tacactccca gttcgaattg cctgccaatg 1980
aaagtggaga aactcaagct atccgactat cagacagcac cgccactcac cgcaacagcc 2040
gcaccagtgc acgattttaa caaaattagc acatacaata tatccgctga gaaatggaga 2100
gaacagccca gcagaagcga ggaaaagaag tcgaactcgc cattgaatga caacaccttt 2160
agctcgaaac ccacaaatgt cgagagtaca agcaaaagcc atgatgttca ttcagcaaat 2220
caaattgatt gcgagaaagt gtgcgcgcag agcagcgata agctaaataa tcatctggcc 2280
gacaagattg tcgagaacaa caatttgat ataggcgggc atgaggaaaa gaagttggtt 2340
cattcgataa gcagcgaaga ctacacacaa atcaaggaca aatcgaatga tttcacaaaa 2400
tttaacgaag ccggctacaa aattctgcaa attaaaagcg acagctcact catctcatcg 2460
aagctatacc aaaagggat acacaaggat aacttgagc gttcgcagag acttacagag 2520
agtgtgaata cgattcccg taatgccacc gccaccgcyg tgagcagcag ctactcacc 2580
aaattcaata taaattcagc aaagccagcc gccgccgcyg attcgcgtag cactggcaca 2640
gatccaagta caccacagaa cattctacag attaaagatt tgaatttccc ctcaaggtcg 2700
tcgtctcgca tatccagcc ggagctgcac tacgccagcc tagatcttcc ccattgcagt 2760
ggccaaaatc cagctaaata cctgaagaga ggatcacgcg aatcgccgcy ggtgtccgca 2820
tgcccggagg atgggaatac ctatgcgaaa atcgactttg accaatccga ctctcttccc 2880
tcctcatcga acatatttaa tacgtaa 2907

<210> 2
 <211> 2907
 <212> DNA
 <213> Drosophila melanogaster

<220>
 <221> CDS
 <222> (1)..(2907)

<400> 2
 atg gca tca ata tcg gat gac ggc atg gcg ctg agt ggc tac ctc aag 48
 Met Ala Ser Ile Ser Asp Asp Gly Met Ala Leu Ser Gly Tyr Leu Lys
 1 5 10 15
 aag ctg aag acc atg aag aag aag ttc ttt gtg ctg tac gag gag acg 96
 Lys Leu Lys Thr Met Lys Lys Lys Phe Phe Val Leu Tyr Glu Glu Thr
 20 25 30
 agc act tcg gca gcc cgg ctg gag tac tac gat acc gaa aag aag ttc 144
 Ser Thr Ser Ala Ala Arg Leu Glu Tyr Tyr Asp Thr Glu Lys Lys Phe
 35 40 45
 ctg caa aga gcc gag cca aaa agg gtt ata tat ctg aag aat tgc ttc 192
 Leu Gln Arg Ala Glu Pro Lys Arg Val Ile Tyr Leu Lys Asn Cys Phe
 50 55 60
 aac atc aat cgc cgt ttg gac acc aag cat aga ttt gtc att gtg ctc 240
 Asn Ile Asn Arg Arg Leu Asp Thr Lys His Arg Phe Val Ile Val Leu
 65 70 75 80
 tcc tcc aga gac ggt gga ttc ggc atc gtt ctc gag aac gaa aat gat 288
 Ser Ser Arg Asp Gly Gly Phe Gly Ile Val Leu Glu Asn Glu Asn Asp
 85 90 95
 tta cgc aaa tgg ttg gac aaa cta cta gtt cta caa agg aac ata gcc 336
 Leu Arg Lys Trp Leu Asp Lys Leu Leu Val Leu Gln Arg Asn Ile Ala
 100 105 110
 aat tcg aat gga aca gcg cac tca cct tat gac cac gtt tgg caa gtt 384
 Asn Ser Asn Gly Thr Ala His Ser Pro Tyr Asp His Val Trp Gln Val
 115 120 125
 gtc att caa aag aag ggt att tcg gag aaa gtt gga atc acc gga acc 432
 Val Ile Gln Lys Lys Gly Ile Ser Glu Lys Val Gly Ile Thr Gly Thr
 130 135 140
 tac cac tgt tgc ctt act tca aaa tcc ctg aca ttc gtg tgc att gga 480
 Tyr His Cys Cys Leu Thr Ser Lys Ser Leu Thr Phe Val Cys Ile Gly
 145 150 155 160
 ccg gag aag acg ccc aat ggc gag gat cgc gtt gcg agc att gaa ata 528
 Pro Glu Lys Thr Pro Asn Gly Glu Asp Arg Val Ala Ser Ile Glu Ile
 165 170 175
 ctt ttg acc acg atc agg cga tgc ggt cat gca tcc cca caa tgt ata 576
 Leu Leu Thr Thr Ile Arg Arg Cys Gly His Ala Ser Pro Gln Cys Ile
 180 185 190
 ttc tac gtg gaa ctt ggc cgc caa agt gtc ttg gga tct ggt gat ctg 624
 Phe Tyr Val Glu Leu Gly Arg Gln Ser Val Leu Gly Ser Gly Asp Leu
 195 200 205

tgg atg gag acg gat aac gca gct att gct act aat atg cac aac acg Trp Met Glu Thr Asp Asn Ala Ala Ile Ala Thr Asn Met His Asn Thr 210 215 220	672
ata ctg agc gct atg tca gcc aaa aca gag tcg aac acg aat tta ata Ile Leu Ser Ala Met Ser Ala Lys Thr Glu Ser Asn Thr Asn Leu Ile 225 230 235 240	720
aac gtt tat cag aat aga cct gac tta agt cac gag ccc atg aga aag Asn Val Tyr Gln Asn Arg Pro Asp Leu Ser His Glu Pro Met Arg Lys 245 250 255	768
cga tcg tcg tct gca aac gaa gca tcg aag ccg ata aac gta aat gtc Arg Ser Ser Ser Ala Asn Glu Ala Ser Lys Pro Ile Asn Val Asn Val 260 265 270	816
ata caa aat agt caa aac tct ctc gaa ttg cgc agc tgc agt tcg ccc Ile Gln Asn Ser Gln Asn Ser Leu Glu Leu Arg Ser Cys Ser Ser Pro 275 280 285	864
cat aac tat ggt ttc ggc aga gag aga tgc gat agc tta cca acc aga His Asn Tyr Gly Phe Gly Arg Glu Arg Cys Asp Ser Leu Pro Thr Arg 290 295 300	912
aat gga acc cta agc gag tcc agc aat caa acg tac ttt ggt tcc aac Asn Gly Thr Leu Ser Glu Ser Ser Asn Gln Thr Tyr Phe Gly Ser Asn 305 310 315 320	960
cat gga ctg cga tcc aat act ata tct ggc atc cgt ccg cac tca acc His Gly Leu Arg Ser Asn Thr Ile Ser Gly Ile Arg Pro His Ser Thr 325 330 335	1008
aac aag cat agt aat agt cca acg ttc acc atg cca tta aga tgc tca Asn Lys His Ser Asn Ser Pro Thr Phe Thr Met Pro Leu Arg Cys Ser 340 345 350	1056
gaa tcc gaa gag tca tca att agt gtc gat gaa tcc gac gac aac ggc Glu Ser Glu Glu Ser Ser Ile Ser Val Asp Glu Ser Asp Asp Asn Gly 355 360 365	1104
agt ttt agc cac tac aga tta aac acg cgg tca tct gag acg gca att Ser Phe Ser His Tyr Arg Leu Asn Thr Arg Ser Ser Glu Thr Ala Ile 370 375 380	1152
cct gag gaa aac att gat gac ttt gcc agt gcg gaa tta ttt agc aaa Pro Glu Glu Asn Ile Asp Asp Phe Ala Ser Ala Glu Leu Phe Ser Lys 385 390 395 400	1200
gtc acc gaa caa aat gta agt gac gaa aac tac ata ccg atg aat cca Val Thr Glu Gln Asn Val Ser Asp Glu Asn Tyr Ile Pro Met Asn Pro 405 410 415	1248
gtc aat cct acc gat gct atc cat gaa aag gag aag gct gat atg cag Val Asn Pro Thr Asp Ala Ile His Glu Lys Glu Lys Ala Asp Met Gln 420 425 430	1296
aga ttg gaa gat gct tcg ctg cat ttc aac ttt ccg gag cac gcg tcg Arg Leu Glu Asp Ala Ser Leu His Phe Asn Phe Pro Glu His Ala Ser 435 440 445	1344
gaa aag ctt gct aag gat ttt gat ctg gac tct gat aac caa tgc tgt Glu Lys Leu Ala Lys Asp Phe Asp Leu Asp Ser Asp Asn Gln Cys Cys 450 455 460	1392

cgt ccc att cgc gcc tat tcg ata ggc aac aag gtt gag cat tta aag	1440
Arg Pro Ile Arg Ala Tyr Ser Ile Gly Asn Lys Val Glu His Leu Lys	
465 470 475 480	
ttt aat aag cgc ctg gga cac ttg aat gat acg gga cag aat ccg aat	1488
Phe Asn Lys Arg Leu Gly His Leu Asn Asp Thr Gly Gln Asn Pro Asn	
485 490 495	
cgc gtg cga gcc tac tcg gtt ggc tcc aaa tcg aag ata ccg cgc tgc	1536
Arg Val Arg Ala Tyr Ser Val Gly Ser Lys Ser Lys Ile Pro Arg Cys	
500 505 510	
gac ctg cag cga gtg gtc ctc gtg gag gac aat aaa cat gag ttc aca	1584
Asp Leu Gln Arg Val Val Leu Val Glu Asp Asn Lys His Glu Phe Thr	
515 520 525	
gcg aat agg agt cag agt agc att acc aag gaa gga acc agc tat ggc	1632
Ala Asn Arg Ser Gln Ser Ser Ile Thr Lys Glu Gly Thr Ser Tyr Gly	
530 535 540	
agc agt gcc aat cga caa aag aag tcc aca agt gct cca ctc ctc agt	1680
Ser Ser Ala Asn Arg Gln Lys Lys Ser Thr Ser Ala Pro Leu Leu Ser	
545 550 555 560	
ctg aag aac cag ata aac tcc gac cga atg agt gac tta atg gaa att	1728
Leu Lys Asn Gln Ile Asn Ser Asp Arg Met Ser Asp Leu Met Glu Ile	
565 570 575	
gat ttt tca caa gca acc aat ttg gaa aag cag aag ttc atc aag aat	1776
Asp Phe Ser Gln Ala Thr Asn Leu Glu Lys Gln Lys Phe Ile Lys Asn	
580 585 590	
aat gaa att ccg aaa tac att gaa aac gtg ttc cca aaa gcc ccg cga	1824
Asn Glu Ile Pro Lys Tyr Ile Glu Asn Val Phe Pro Lys Ala Pro Arg	
595 600 605	
acg gat agc tcc agc cta act ctg cac gcc aca agt caa aag gac att	1872
Thr Asp Ser Ser Ser Leu Thr Leu His Ala Thr Ser Gln Lys Asp Ile	
610 615 620	
ttc aat ggc acc aaa cta aat aac act gcg atc aca tcc gag gat ggt	1920
Phe Asn Gly Thr Lys Leu Asn Asn Thr Ala Ile Thr Ser Glu Asp Gly	
625 630 635 640	
tac ctc gag atg aag cca gtc ggt aat gga tac act ccc agt tcg aat	1968
Tyr Leu Glu Met Lys Pro Val Gly Asn Gly Tyr Thr Pro Ser Ser Asn	
645 650 655	
tgc ctg cca atg aaa gtg gag aaa ctc aag cta tcc gac tat cag aca	2016
Cys Leu Pro Met Lys Val Glu Lys Leu Lys Leu Ser Asp Tyr Gln Thr	
660 665 670	
gca ccg cca ctc acc gca aca gcc gca cca gtg cac gat tta aac aaa	2064
Ala Pro Pro Leu Thr Ala Thr Ala Ala Pro Val His Asp Leu Asn Lys	
675 680 685	
att agc aca tac aat ata tcc gct gag aaa tgg aga gaa cag ccc agc	2112
Ile Ser Thr Tyr Asn Ile Ser Ala Glu Lys Trp Arg Glu Gln Pro Ser	
690 695 700	
aga agc gag gaa aag aag tcg aac tcg cca ttg aat gac aac acc ttt	2160
Arg Ser Glu Glu Lys Lys Ser Asn Ser Pro Leu Asn Asp Asn Thr Phe	
705 710 715 720	

agc tcg aaa ccc aca aat gtc gag agt aca agc aaa agc cat gat gtt	2208
Ser Ser Lys Pro Thr Asn Val Glu Ser Thr Ser Lys Ser His Asp Val	
725 730 735	
cat tca gca aat caa att gat tgc gag aaa gtg tgc gcg cag agc agc	2256
His Ser Ala Asn Gln Ile Asp Cys Glu Lys Val Cys Ala Gln Ser Ser	
740 745 750	
gat aag cta aat aat cat ctg gcc gac aag att gtc gag aac aac aat	2304
Asp Lys Leu Asn Asn His Leu Ala Asp Lys Ile Val Glu Asn Asn Asn	
755 760 765	
ttg gat ata ggc ggg cat gag gaa aag aag ttg gtt cat tcg ata agc	2352
Leu Asp Ile Gly Gly His Glu Glu Lys Lys Leu Val His Ser Ile Ser	
770 775 780	
agc gaa gac tac aca caa atc aag gac aaa tcg aat gat ttc aca aaa	2400
Ser Glu Asp Tyr Thr Gln Ile Lys Asp Lys Ser Asn Asp Phe Thr Lys	
785 790 795 800	
ttt aac gaa gcc ggc tac aaa att ctg caa att aaa agc gac agc tca	2448
Phe Asn Glu Ala Gly Tyr Lys Ile Leu Gln Ile Lys Ser Asp Ser Ser	
805 810 815	
ctc atc tca tcg aag cta tac caa aag ggt ata cac aag gat aac ttg	2496
Leu Ile Ser Ser Lys Leu Tyr Gln Lys Gly Ile His Lys Asp Asn Leu	
820 825 830	
gag cgt tcg cag aga ctt aca gag agt gtg aat acg att ccc gat aat	2544
Glu Arg Ser Gln Arg Leu Thr Glu Ser Val Asn Thr Ile Pro Asp Asn	
835 840 845	
gcc acc gcc acc gcg gtg agc agc agc tca ctc acc aaa ttc aat ata	2592
Ala Thr Ala Thr Ala Val Ser Ser Ser Ser Leu Thr Lys Phe Asn Ile	
850 855 860	
aat tca gca aag cca gcc gcc gcc gcc gat tcg cgt agc act ggc aca	2640
Asn Ser Ala Lys Pro Ala Ala Ala Asp Ser Arg Ser Thr Gly Thr	
865 870 875 880	
gat cca agt aca cca cag aac att cta cag att aaa gat ttg aat ttc	2688
Asp Pro Ser Thr Pro Gln Asn Ile Leu Gln Ile Lys Asp Leu Asn Phe	
885 890 895	
ccc tca agg tcg tcg tct cgc ata tcc cag ccg gag ctg cac tac gcc	2736
Pro Ser Arg Ser Ser Ser Arg Ile Ser Gln Pro Glu Leu His Tyr Ala	
900 905 910	
agc cta gat ctt ccc cat tgc agt ggc caa aat cca gct aaa tac ctg	2784
Ser Leu Asp Leu Pro His Cys Ser Gly Gln Asn Pro Ala Lys Tyr Leu	
915 920 925	
aag aga gga tca cgc gaa tcg ccg ccg gtg tcc gca tgc ccg gag gat	2832
Lys Arg Gly Ser Arg Glu Ser Pro Pro Val Ser Ala Cys Pro Glu Asp	
930 935 940	
ggg aat acc tat gcg aaa atc gac ttt gac caa tcc gac tcc tct tcc	2880
Gly Asn Thr Tyr Ala Lys Ile Asp Phe Asp Gln Ser Asp Ser Ser Ser	
945 950 955 960	
tcc tca tcg aac ata ttt aat acg taa	2907
Ser Ser Ser Asn Ile Phe Asn Thr	
965	

<210> 3
 <211> 968
 <212> PRT
 <213> Drosophila melanogaster

<400> 3

Met	Ala	Ser	Ile	Ser	Asp	Asp	Gly	Met	Ala	Leu	Ser	Gly	Tyr	Leu	Lys
1				5					10					15	
Lys	Leu	Lys	Thr	Met	Lys	Lys	Lys	Phe	Phe	Val	Leu	Tyr	Glu	Glu	Thr
			20					25					30		
Ser	Thr	Ser	Ala	Ala	Arg	Leu	Glu	Tyr	Tyr	Asp	Thr	Glu	Lys	Lys	Phe
		35					40					45			
Leu	Gln	Arg	Ala	Glu	Pro	Lys	Arg	Val	Ile	Tyr	Leu	Lys	Asn	Cys	Phe
	50					55					60				
Asn	Ile	Asn	Arg	Arg	Leu	Asp	Thr	Lys	His	Arg	Phe	Val	Ile	Val	Leu
65					70					75					80
Ser	Ser	Arg	Asp	Gly	Gly	Phe	Gly	Ile	Val	Leu	Glu	Asn	Glu	Asn	Asp
				85					90					95	
Leu	Arg	Lys	Trp	Leu	Asp	Lys	Leu	Leu	Val	Leu	Gln	Arg	Asn	Ile	Ala
			100					105					110		
Asn	Ser	Asn	Gly	Thr	Ala	His	Ser	Pro	Tyr	Asp	His	Val	Trp	Gln	Val
		115					120					125			
Val	Ile	Gln	Lys	Lys	Gly	Ile	Ser	Glu	Lys	Val	Gly	Ile	Thr	Gly	Thr
	130					135					140				
Tyr	His	Cys	Cys	Leu	Thr	Ser	Lys	Ser	Leu	Thr	Phe	Val	Cys	Ile	Gly
145					150					155					160
Pro	Glu	Lys	Thr	Pro	Asn	Gly	Glu	Asp	Arg	Val	Ala	Ser	Ile	Glu	Ile
				165					170					175	
Leu	Leu	Thr	Thr	Ile	Arg	Arg	Cys	Gly	His	Ala	Ser	Pro	Gln	Cys	Ile
			180					185					190		
Phe	Tyr	Val	Glu	Leu	Gly	Arg	Gln	Ser	Val	Leu	Gly	Ser	Gly	Asp	Leu
		195					200					205			
Trp	Met	Glu	Thr	Asp	Asn	Ala	Ala	Ile	Ala	Thr	Asn	Met	His	Asn	Thr
	210					215					220				
Ile	Leu	Ser	Ala	Met	Ser	Ala	Lys	Thr	Glu	Ser	Asn	Thr	Asn	Leu	Ile
225					230					235					240
Asn	Val	Tyr	Gln	Asn	Arg	Pro	Asp	Leu	Ser	His	Glu	Pro	Met	Arg	Lys
				245					250					255	
Arg	Ser	Ser	Ser	Ala	Asn	Glu	Ala	Ser	Lys	Pro	Ile	Asn	Val	Asn	Val
			260					265					270		
Ile	Gln	Asn	Ser	Gln	Asn	Ser	Leu	Glu	Leu	Arg	Ser	Cys	Ser	Ser	Pro
	275						280					285			
His	Asn	Tyr	Gly	Phe	Gly	Arg	Glu	Arg	Cys	Asp	Ser	Leu	Pro	Thr	Arg
	290					295					300				

Asn	Gly	Thr	Leu	Ser	Glu	Ser	Ser	Asn	Gln	Thr	Tyr	Phe	Gly	Ser	Asn	
305					310					315					320	
His	Gly	Leu	Arg	Ser	Asn	Thr	Ile	Ser	Gly	Ile	Arg	Pro	His	Ser	Thr	
				325					330					335		
Asn	Lys	His	Ser	Asn	Ser	Pro	Thr	Phe	Thr	Met	Pro	Leu	Arg	Cys	Ser	
			340					345					350			
Glu	Ser	Glu	Glu	Ser	Ser	Ile	Ser	Val	Asp	Glu	Ser	Asp	Asp	Asn	Gly	
		355					360					365				
Ser	Phe	Ser	His	Tyr	Arg	Leu	Asn	Thr	Arg	Ser	Ser	Glu	Thr	Ala	Ile	
	370					375					380					
Pro	Glu	Glu	Asn	Ile	Asp	Asp	Phe	Ala	Ser	Ala	Glu	Leu	Phe	Ser	Lys	
385					390					395					400	
Val	Thr	Glu	Gln	Asn	Val	Ser	Asp	Glu	Asn	Tyr	Ile	Pro	Met	Asn	Pro	
				405					410					415		
Val	Asn	Pro	Thr	Asp	Ala	Ile	His	Glu	Lys	Glu	Lys	Ala	Asp	Met	Gln	
			420					425					430			
Arg	Leu	Glu	Asp	Ala	Ser	Leu	His	Phe	Asn	Phe	Pro	Glu	His	Ala	Ser	
		435					440					445				
Glu	Lys	Leu	Ala	Lys	Asp	Phe	Asp	Leu	Asp	Ser	Asp	Asn	Gln	Cys	Cys	
	450					455					460					
Arg	Pro	Ile	Arg	Ala	Tyr	Ser	Ile	Gly	Asn	Lys	Val	Glu	His	Leu	Lys	
465					470					475					480	
Phe	Asn	Lys	Arg	Leu	Gly	His	Leu	Asn	Asp	Thr	Gly	Gln	Asn	Pro	Asn	
				485					490					495		
Arg	Val	Arg	Ala	Tyr	Ser	Val	Gly	Ser	Lys	Ser	Lys	Ile	Pro	Arg	Cys	
			500					505					510			
Asp	Leu	Gln	Arg	Val	Val	Leu	Val	Glu	Asp	Asn	Lys	His	Glu	Phe	Thr	
		515					520					525				
Ala	Asn	Arg	Ser	Gln	Ser	Ser	Ile	Thr	Lys	Glu	Gly	Thr	Ser	Tyr	Gly	
	530					535					540					
Ser	Ser	Ala	Asn	Arg	Gln	Lys	Lys	Ser	Thr	Ser	Ala	Pro	Leu	Leu	Ser	
545					550					555					560	
Leu	Lys	Asn	Gln	Ile	Asn	Ser	Asp	Arg	Met	Ser	Asp	Leu	Met	Glu	Ile	
				565					570					575		
Asp	Phe	Ser	Gln	Ala	Thr	Asn	Leu	Glu	Lys	Gln	Lys	Phe	Ile	Lys	Asn	
			580					585					590			
Asn	Glu	Ile	Pro	Lys	Tyr	Ile	Glu	Asn	Val	Phe	Pro	Lys	Ala	Pro	Arg	
	595						600					605				
Thr	Asp	Ser	Ser	Ser	Leu	Thr	Leu	His	Ala	Thr	Ser	Gln	Lys	Asp	Ile	
	610					615					620					
Phe	Asn	Gly	Thr	Lys	Leu	Asn	Asn	Thr	Ala	Ile	Thr	Ser	Glu	Asp	Gly	
625					630					635					640	

Tyr	Leu	Glu	Met	Lys	Pro	Val	Gly	Asn	Gly	Tyr	Thr	Pro	Ser	Ser	Asn
				645					650					655	
Cys	Leu	Pro	Met	Lys	Val	Glu	Lys	Leu	Lys	Leu	Ser	Asp	Tyr	Gln	Thr
			660					665					670		
Ala	Pro	Pro	Leu	Thr	Ala	Thr	Ala	Ala	Pro	Val	His	Asp	Leu	Asn	Lys
			675				680					685			
Ile	Ser	Thr	Tyr	Asn	Ile	Ser	Ala	Glu	Lys	Trp	Arg	Glu	Gln	Pro	Ser
	690					695					700				
Arg	Ser	Glu	Glu	Lys	Lys	Ser	Asn	Ser	Pro	Leu	Asn	Asp	Asn	Thr	Phe
705					710					715					720
Ser	Ser	Lys	Pro	Thr	Asn	Val	Glu	Ser	Thr	Ser	Lys	Ser	His	Asp	Val
				725					730					735	
His	Ser	Ala	Asn	Gln	Ile	Asp	Cys	Glu	Lys	Val	Cys	Ala	Gln	Ser	Ser
			740					745					750		
Asp	Lys	Leu	Asn	Asn	His	Leu	Ala	Asp	Lys	Ile	Val	Glu	Asn	Asn	Asn
		755					760					765			
Leu	Asp	Ile	Gly	Gly	His	Glu	Glu	Lys	Lys	Leu	Val	His	Ser	Ile	Ser
	770					775					780				
Ser	Glu	Asp	Tyr	Thr	Gln	Ile	Lys	Asp	Lys	Ser	Asn	Asp	Phe	Thr	Lys
785					790					795					800
Phe	Asn	Glu	Ala	Gly	Tyr	Lys	Ile	Leu	Gln	Ile	Lys	Ser	Asp	Ser	Ser
				805					810					815	
Leu	Ile	Ser	Ser	Lys	Leu	Tyr	Gln	Lys	Gly	Ile	His	Lys	Asp	Asn	Leu
			820					825					830		
Glu	Arg	Ser	Gln	Arg	Leu	Thr	Glu	Ser	Val	Asn	Thr	Ile	Pro	Asp	Asn
		835					840					845			
Ala	Thr	Ala	Thr	Ala	Val	Ser	Ser	Ser	Ser	Leu	Thr	Lys	Phe	Asn	Ile
	850					855					860				
Asn	Ser	Ala	Lys	Pro	Ala	Ala	Ala	Ala	Asp	Ser	Arg	Ser	Thr	Gly	Thr
865					870					875					880
Asp	Pro	Ser	Thr	Pro	Gln	Asn	Ile	Leu	Gln	Ile	Lys	Asp	Leu	Asn	Phe
				885					890					895	
Pro	Ser	Arg	Ser	Ser	Ser	Arg	Ile	Ser	Gln	Pro	Glu	Leu	His	Tyr	Ala
			900					905					910		
Ser	Leu	Asp	Leu	Pro	His	Cys	Ser	Gly	Gln	Asn	Pro	Ala	Lys	Tyr	Leu
		915						920				925			
Lys	Arg	Gly	Ser	Arg	Glu	Ser	Pro	Pro	Val	Ser	Ala	Cys	Pro	Glu	Asp
	930					935					940				
Gly	Asn	Thr	Tyr	Ala	Lys	Ile	Asp	Phe	Asp	Gln	Ser	Asp	Ser	Ser	Ser
945					950					955					960
Ser	Ser	Ser	Asn	Ile	Phe	Asn	Thr								
				965											

<210> 4
<211> 5210
<212> DNA
<213> *Drosophila melanogaster*

<400> 4
agaacgactt tttcctcctt agtcagtcac aagaaaacta aagcttacca acaatacggc 60
gtgtattgtt aaattattac aacaaataaa atattcaaat tgtatttaaa aatatagtaa 120
ccattaaaaa ataaaatcaa tatgcgaaac tttgtaattt cttactcacc cttgtttttt 180
gagcccgctt tcttaagtta aatcggttaa ataccagttt aatcatttca ttgtcctgat 240
ttcaggagct aattacattt taatctttgt ataaattcat aaaattaaaa tggaaatgtt 300
taaccacata aaatatttgg gtatataaag tcgatacata ctttttaaaat tttgtttcat 360
acaagaatat ggaaagtaga taatttagtt accgattaaa acattttctaa aatacaaaaa 420
atttaaaatg atgattaaat aaaaaactgt tataactaaat ttaaacgaaa caaacggtca 480
ttcgataact caattagtat cgaataagcc ggcgtgtaaa tcgggttgcc aactctcacc 540
ggtgtagaga tcgggatggc aacttcgtat tgttattcct atgctgcgat aacgataaca 600
gcaaagtgtc atcggaata aatgtgggtt tgataataaa gaattgcgtt ggtgaacaat 660
aataatacaa atgctaaatg tatcgcgagg ataactagta aacactgatt tcgcgcacat 720
cgggcatagc ggcagctaga cgtcttaggt aacacatttc ccagccacat tggcgttgag 780
gtattatttc ccatatccat gtgcgtttgt aaatgatacc accagagtgt gccatatgta 840
tcgttgtttg tacatacatg cctacgggga aaaataactc gcagatacat atgtatgtaa 900
gatgtattag aattccacag atatatgtac atataaatat atatatatat ttatgatgct 960
catctataaa aaaaaatag tacaacata cgcgcgcacg tacttatgta tgtacatata 1020
catatataca tatatgtata ataaaatgag catctagctg cggttatctt aatgcaatgc 1080
gcagaaacct gaaaacgaaa taaacaatct ttacagcgcc agcacagtga gccaattatg 1140
aattcacaat tccacatcca attccgattc cgaattccat cgctttacat cctaattcga 1200
atacgtcgcg ccgcgtaagc tgcactcgaa tattgacata aacgacgtaa ttgcgtgttt 1260
tgattgcgat ttccgatgct cgatgttgac agacggcaag gatttttttt gccagccgac 1320
attgcgaatg cttttcgcgt gtttgcggtt tgaaaagcga tttgttggtc cggcagttgg 1380
aaatgttttg tgctgtttat gcttagaaaa gcaaatgtat tgcacaaaac tcgtgtttcc 1440
tacattagta aagcctataa cttagggtata tgtttctaaa ttacaatgca aaaaataaaa 1500
aacattatac atgtgttcgt tctttaattt gaaaacagaa aagtgaagc cttgcaaatc 1560
aaatatgtgt ccatatcgcc tactaataat ataaacagc cgccttcag gaactaagat 1620
tggaaatcat ggcatcaata tcggatgacg gcatggcgct gagtggctac ctcaagaagc 1680
tgaagaccat gaagaagaag ttctttgtgc tgtacgagga gacgagcact tcggcagccc 1740

ggctggagta ctacgatacc gaaaagaagt tcctgcaaag agccgagcca aaaaggggta 1800
 tctatctgaa gaattgcttc aacatcaatc gccgtttgga caccaagcat agatttgtca 1860
 ttgtgctctc ctccagagac ggtggattcg gcacggttct cgagaacgaa aatgatttac 1920
 gcaaatgggtt ggacaaacta ctagtcttac aaaggaacat agccaattcg aatggaacag 1980
 cgcactcacc ttatgggtatg ccaaataaac tataactacc agttagtatg aaacctaaac 2040
 acttcatttc actttgcaga ccacgtttgg caagttgtca ttcaaaagaa gggattttcg 2100
 gagaaagttg gaatcaccgg aacctaccac tgttgcccta cttcaaaatc cctgacattc 2160
 gtgtgcattg gaccggagaa gacgcccaat ggcgaggatc gcgttgcgag cattgaaata 2220
 cttttgacca cgatcaggcg gttagtgtt gccagcaaaa ctgcaaggga ttgtaaaata 2280
 attcggactt aatttcagat gcggtcatgc atccccacaa tgtatattct acgtggaact 2340
 tggccgcca agtgtcttgg gatctggtga tctgtggatg gagacggata acgcagctat 2400
 tgctactaat atgcacaaca cgatactgag gtatttagct ctcatataca ctaatccaag 2460
 atttcagat catcctacaa aacgacatag atagtttaag atatctccca gttaacttta 2520
 ataattctgt gggttttttc tttcagcgct atgtcagcca aaacagagtc gaacacgaat 2580
 ttaataaacg tttatcagaa cagacctgac ttaagtcacg agcccatgag aaagcgatcg 2640
 tcgtctgcaa acgaagcatc gaagccgata aacgtaaagtc tcatacaaaa tagtcaaaac 2700
 tctctcgaat tgccgagctg cagttcgccc cataactatg gtaaatactt caaatgtatg 2760
 tttaaacgca aaattaatca aacgcaatcg tttcaggttt cggcagagag agatgcgata 2820
 gcttaccac cagaaatgga accctaagcg agtccagcaa tcaaacgtac tttggttcca 2880
 accatggact gcgatccaat actatatctg gcacccgtcc gcactcaacc aacaagcata 2940
 gtaatagtcc aacgttcacc atgccattaa gatgctcaga atccgaagag tcatcaatta 3000
 gtgtcgatga atccgacgac aacggcagtt ttagccacta cagattaaag tgcgttgcta 3060
 tcaaataata attatttaat aataatcacc atttcaattt ctagcacgcg gtcactctgag 3120
 acggcaattc ctgaggaaaa cattgatgac tttgccagtg cggaattatt tagcaaagtc 3180
 accgaacaaa atggttaagcc aaacacaaaa acaatttttt aacatgaaaa gtagctaate 3240
 aattggcttt gtttaactgc agtaagtgc gaaaactaca taccgatgaa tccagtcaat 3300
 cctaccgatg ctatccatga aaaggagaag gctgatatgc agagattgga agatgcttcg 3360
 ctgcatttca actttccgga gcacgcgctg gaaaagcttg ctaaggattt tgatctggac 3420
 tctgataacc agtgagtaca catttcgctt caactgtgcc acgtaatgca atcaatcaca 3480
 tcttgttaca gatgctgtcg tccatttcgc gcctattcga taggcaacaa ggttgagcat 3540
 ttaaagttta ataagcgctt gggacacttg aatgatacgg gacagaatcc gaatcgctg 3600
 cgagcctact cggttggctc caaatcgaag ataccgcgct gcgacctgca gcgagtggct 3660

ctctgtggagg acaataaaca tgagttcaca gcgaatagga gtcagagtag cattaccaag 3720
 gaaggaacca gctatggcag cagtgccaat cgacaaaaga agtccacaag tgctccactc 3780
 ctctgtctga agaaccagat aaactccgac cgaatgagtg acttaatgga aattgatttt 3840
 tcacaagcaa ccaatttggg aaagcagaag ttcatacaga ataataaata tccgaaatac 3900
 attgaaaacg tgttcccaaa agccccgcga acggatagct ccagcctaac tctgcacgcc 3960
 acaagtcaaa aggacatttt caatggcacc aaactaaata acactgcgat cacatccgag 4020
 gatgggttacc tcgagatgaa gccagtcggt aatggatata ctcccagttc gaattgcctg 4080
 ccaatgaaag tggagaggct caagctatcc gactatcaga cagcaccgcc aatcaccgca 4140
 acagccgcac cagtgcacga tttaaacaaa attagcacat acaatatatc cgctgagaaa 4200
 tggagagAAC agcccagcag aagcgaggaa aagaagtcca actcgccatt gaatgacaac 4260
 acctttggct tgaaaccac aaatgtcgag agtacaagca aaagccatga tgttcattca 4320
 gcaaatcaaa ttgattccga gaaagtgtgc gcgcagagca gcgataagct aaataatctg 4380
 gccgacaaga ttgtcgagaa caacaatttg gatataggcg ggcatgagga aaagaagttg 4440
 gttcattcga taagcagcga agactacaca caaatcaagg acaaatcgaa tgatttcaca 4500
 aaattttaacg aagccggcta caaaattctg caaattaaaa gcgacagctc actcatctca 4560
 tcgaagctat accaaaaggg tatacacaag gataacttgg agcgttcgca gagacttacg 4620
 gagagtgtga atacgattcc cgataatgcc accgccaccg cggtagagcag cagctcactc 4680
 accaaattca atataaattc agcaaagcca gccgccgccg ccgattcgcg tagcactggc 4740
 acagatccaa gtacaccaca gaacattcta cagattaaag atttgaattt cccctcaagg 4800
 tcgtcgtctc gcatatccca gccggagctg cactacgcca gcctagatct tccccattgc 4860
 agtggccaaa atccagctaa atacctgaag agaggatcac gcgaatcgcc gccggtgtcc 4920
 gcatgcccgg aggatgggaa tacctatgcg aaaatcgact ttgaccaatc cgactcctct 4980
 tcctcctcat cgaacatatt taatacgtaa agttttgaaa tttatgacct taccctatat 5040
 atatgatttg tttaatatg tacatttatt gtaaataatc tctgacaagc aaagcttaca 5100
 attttggatg ctaataaata aattttattt aaattataat gatccctttg gacttttttt 5160
 ttttttggac taagaaatca ctactaaaga agggcttttc gaggggttaaa 5210